PROCESS FOR THE EVOLUTIVE DESIGN AND SYNTHESIS OF FUNCTIONAL POLYMERS BASED ON DESIGNER ELEMENTS AND CODES

Publication number: WO9517413
Publication date: 1995-06-29
Inventor: SCHWIENHORST ANDREAS (DE); LINDEMANN BJOERN (DE); FUCHS MERLE (DE); HENCO KARSTEN (DE)
Applicant: EVOTEC BIOSYSTEMS GMBH (DE); SCHWIENHORST ANDREAS (DE); LINDEMANN BJOERN (DE); FUCHS MERLE (DE); HENCO KARSTEN (DE)
Classification:
- international: ***C07H21/00; C07K1/04; C12P19/34;*** **C07H21/00; C07K1/00; C12P19/00;** (IPC1-7): C07H21/00; C07K1/04; C12N15/10; C12P21/00
- European: C07H21/00C4; C07K1/04C; C12P19/34
Application number: WO1994EP04240 19941220
Priority number(s): DE19934343591 19931221

Also published as:
DE4343591 (A1)

Cited documents:
WO9000626
WO9218645

Abstract of **WO9517413**

A process for the production of oligomeric or polymeric functional elements from designer elements in which the functional elements are obtainable by the linking of at least two designer elements, at least one of which is itself made up of at least two monomers which are linked by at least one chemical bond which corresponds to the chemical bond between two designer elements.

Description of DE4343591	Print	Copy	Contact Us	Close

Result Page

Subject-matter of the present invention is a method according to claim 1.

The rapid development of the last years within the field of life sciences does not only have the basic research, but straight also the applied research in this field stimulates. Proteins play here due to its broad action spectrum e outstanding roller. A whole branch of the modern biotechnology is occupied today with that so-called. ?Protein engineering?, D. h. the preparation of designer proteins, which are developed either on the basis of well-known proteins by gradual amending or by complete new synthesis. One differentiates between here above all two beginnings, the rational and the irrational Design.

Rational Design is out to produce an amino acid sequence which folds into a desired structure, and additionally the hoped for function exhibits. Thus this strategy depends protein completely obviously on a deep understanding ?folding?. Progress in the last years concerned and. A. the rational Design of simple structure domains. The Design of larger proteins with complex or unparalleled new properties lies however still outside of the possibilities of this beginning. In contrast to this irrational Design does not set information about the protein structure, protein folding etc. ahead. Only the knowledge of the desired property and a possibility of evaluating molecule populations based on this property are here a condition on the basis of one ?combinatorial libr from peptides or proteins molecules with the desired property are selected and only afterwards analyzed. Here the mechanism, after a molecule masters the task posed, is thus determined not in ahead.

Although this beginning in very elegant way straight also in recent time peptides with simple and z. T. the problem brought, places themselves also here new properties out, as one can come to larger proteins with more complex functions. Already a complete bank of a 20mers supplies with 20< 2> < 0> = 10< 2> < 6> different sequences an astronomically high number examining molecules. Even if the Peptidsequenz is to be coded still by a nucleic acid, the problem places itself in still serious way. Since the genetic code degenerated, D. h. a amino acid and. And. by several different Codons, arises here a number of at least 4< is represented; 6> < 0> = 10< 3> < 6> Molecules, which are synthesized. Normally will at the third Codonposition only G or C certified, in order to avoid Stopcodons to a large extent. The remaining number of 10< 3> < 0> Molecules exceeds still the standard yield of a commercial DNA synthesis by 12 orders of magnitude. A further reducing of the Codons certified per position was suggested by Youvan. Whether this method does not limit the measurable sequence area in insufficient way, straight with the search for new functions, remains being waiting.

For the structure of functional structures nature with modular systems works. Admit are the nucleotide components, the amino acid components (coded as nucleotide triplet) and Exon domains (composed of amino acid components). The evolutive optimization of functional bio polymers according to the patent application WHERE 92/18645 goes out with the conception, by continuous improvement of existing basis characteristics, z. B. an enzyme to find with the continuous adjustment to desired reaction conditions such as ion strength, temperature, pH value an optimal structure. If favourable or at least neutral mutations are possible, then are by repeated repetitions of selection and mutation also removed ranges of the sequence area accessible, which were not taken off by the output population. From the original, already functional structure one departs however with this procedure in no step. Is optimized a property of the output molecule, which is inherent already - although in modest mass - in the original molecule. The ?path?, which such an evolution takes by the sequence area, is certain by the accessible, toward the optima leading burrs in the underlying value landscape. As is the case for all methods, which open the sequence area not completely, the danger which can be estimated only heavily exists to be stuck in a local optimum with this procedure. For practice this means that certain regions of the sequence area including the optima there present, by broad and deep valleys are separated. With the limited population size of molecule species in experiments (P 43 22 147, WHERE 92/18645) however the probability is too low to produce distant Vielfehlermutanten which are beyond this limit and which to these new optima indicates way.

top

Nature developed a number of mechanisms to deal with this problem: long development periods, recombination procedure (horizontal gene transfer, Crossingover, gene conversion, Exon recombination (exon shuffling), virus shuttle, mobile elements (Transposons), subunit structure of complex proteins)) as well as multi-gene families with pseudo genes.

With the number of the parallel led Mutantenbildung and selection the chance can be increased to production of a desired Vielfehlermutante; by recombination mutated gene segments can be mixed efficiently. Function lots of pseudo genes as members of a functional multi-gene family can be received as Vielfehlermutanten even over longer development periods without Gegenselektion in their existence, in order to become possibly during back preservation of a function again positively selectable.

The transmission of these mechanisms on an efficient in vitro optimization is obviously not so easily possible. The difficulties must be solved however in each case for such setting of tasks, with which a continuous optimization cannot be expected. This applies in particular to such adjustment processes, with which a function must completely again be

established.

That the invention underlying technical problem concerns the supply of a method to the preparation of oligomerer or polymere functional elements such as bio polymers with functional properties, for example enzymes, Ribozymen, active substances, etc. The conventional Screening methods superior a method is to be made available under utilization of evolutiver strategies.

This problem is solved by a method with the features of the claim 1. The Unteransprüche following to it concern themselves preferential embodiments of the method according to invention.

To the preparation of oligomerer or polymere functional elements from form elements first form elements are developed according to invention by chemical or enzymatic linkage of at least two monomers and the so available form elements to functional elements are then linked. Nature corresponds to the chemical connection between the monomers of those between the respective form elements. The so available functional elements can be tested then on the certain potential functions. The advantages of the proceeding according to invention are continued to clarify by the following description.

Preferred the linkage of the form elements under inset of a solid phase is accomplished as poorly reactive. The linkage of the form elements can take place chemically and/or enzymatically. The linkage of the form elements to the functional elements can take place either according to plan via purposeful addition of the individual form elements and following linkage or also statistical via coincidentally controlled addition of the functional elements and their linkage. It is possible to accomplish the linkage gradually constructing stereospezifisch and/or arranged.

As form elements are preferably applicable nucleic acids, doppelsträngige or einzelsträngige DNA and/or RNA and/or modified nucleic acids. As form elements are applicable also peptides and/or Polypeptide and/or other couplingable chemical Oligomer Formenelemente. In addition also Oligo can belong or polysaccharides.

In a preferential embodiment of the method according to invention the form elements are used as Oligomerbausteine already synthesized or manufactured in the reaction container quasi in situ.

It is favourable to accomplish the reaction of the form elements in parallel led micro reaction beginnings (suggested as in P 43 22 147,5) with which the form elements in pre-determined order are linked. In particular after synthesis the reaction products are continued to process such as functional elements or preliminary stages of it bound at the solid phase to remain and after separation of the reaction partners or decoupled by the solid phase. It is however likewise possible, the reaction in suitable, for the person skilled in the art admitted reaction conditions in solution to accomplish or the festphasengekoppelte or in homogeneous solution accomplished reaction to combine.

By inset of the fluorescence correlation spectroscopy (FK) (PCT/EP 93/01291) it is made possible to evaluate the function mode of the functional elements in the same volume element directly in which also the synthesis runs off. This means a very direct possibility of controlling the result of a developing functional element synthesis.

Preferably per reaction step, with the gradual linkage that is coupled form element, in each case a form element as reaction partner at solid phase. Also mixtures can be used by form elements and/or be generated directly in the reaction container. As form elements if nucleic acids are used, then it is favourable to use at least one reaction partner with an interface restriction enzyme too provided or a nucleic acid form element which is free from starting and/or Stopcodons. Preferably the reaction interfaces are such, which can be recognized by restriction enzymes of the class IIS. The introduction of restriction interfaces of this enzyme class is favourable, since arbitrary sequences can be linked arranged, without the choice of the reaction enzyme affects the sequence requirements of the final product.

If form elements einzelsträngige overhangs which can be linked are imported into that, then directed via it arbitrary sequences can be linked, without thereby any demands must be made against the sequence of the desired final product. This requirement can be obtained also by selective phosphorylation in place of and in combination with the introduction of the einzelsträngigen overhangs.

The method according to invention permits the inset of form elements, which admits after X-ray-crystallographically analyzed natural function domains of proteins and Polypeptiden is. So already well-known components can do and/or. Modules by in nature functional elements already occurring to be used.

The form elements which can be used can be won also from selection experiments.

The use of form elements is in particular favourable in a length from 1 to 60 amino acids or nucleotide sequences of appropriate coding length. The form elements can be also at certain positions degeniert and/or Deletionen or Insertionen carry, in particular with use of nucleotides as form elements.

Also the use of the method according to invention is above stressed like described parallel to the synthesis constructed it form libraries functional Oligomere or polymers.

The original task of ?combinatorial LIBRARies? is rather the offer function variety as sequence variety. It is today a fact that the three-dimensional Stru of proteins is relatively stable against substitutions of individual amino acids. By the large number of enlightened protein structures one won the realization that proteins can exhibit no or only very small Sequenzhomologie, but although the same or very similar 3D-Struktur can take. This is based possibly on the fact that only a limited number of possible folding ways of amino acid chains under biological conditions is stable. In addition, structural relationship reflects the evolution of rezenter proteins from a relatively limited number of Ur-structures, - modules out again. These modules can to be understood as small, functional domains or compact structure units and be able also in today's genes to be easily sought out. In the hypothesis ?Exon shuffling? it is assumed that the evolution was thus accelerated to more complex proteins straight by the combination of Exons, modules in the sense described above enormously. If one assumes that the number of the Exons, which the construction of all today well-known proteins would permit between 1000 and 7000 is to be searched, opens a hierarchical strategy to ?protein Design? with components of increasing complexity the possibility of the many faster Durchmessung ?shape space? with associated ? fitness landscape? than it the search in a traditional ?combinatorial LIBRARY? would permit. A protein from 150 amino acids (the value of a classical nucleotide connection place, the so-called. ?Rossman fold?) would have after conventional

method from a library of 20< 1> < 5> < 0> = 10< 1> < 9> < 5> different amino acid sequences to be selected. Combinations of 1000 different modules of the length 30 amino acids result in however only one complexity of 1000< 5> = 10< 1> < 5> Molecules.

The method according to invention is a hierarchical method to the Design of proteins, nucleic acids their derivatives or chemical Oligo or polymers with certain desired properties, on the basis of module - libraries, in the following as form elements designates. According to invention the form elements can be also gene segments, which code for form elements. The form elements functioning as modules should be coincidentally combinable. Smaller proteins or subunits for larger proteins with certain properties are out-separated in a following selection step from the pool by module combinations and can serve for their part again than components in a subunit library, etc.

On each construction stage additionally a ?noise? can be introduced on amino acid sequence level by incorrect copying of individual components. This makes modulating and thus a further functional optimization of selected molecules for the three-dimensional arrangement possible of chemical groups. The suggested strategy requires a new kind of ?Artificial of genes assembly?.

So far above all two methods are used, it is common to which that the DNA in a certain orientation is ligiert, in order to specify thereby also the succession of the amino acids. The probably oldest method - developed by Khorana and its coworkers - works with overlapping complementary einzelsträngigen DNA molecules, which are hybridized with one another before the Ligation. The second method uses interfaces of restriction enzymes in the gene which can be designed, in order to divide in these places the gene into blocks, which are then built up in several successive steps. By both methods the sequence at the transitions of the used Oligo DNAs becomes and/or. Blocks methodically causes fixed. This however does not correspond straight to the requirement already after arbitrary exchangeability of the individual modules in the construction phase a gene component of the present invention is thus necessarily also a new kind of the ?Artificial of genes assembly?. In general form will proceed as follows according to invention:

- The method of the ?Artificial of genes assembly? works the analog in the WHERE 92/18645 described method;
- the method deduces not handling the variance in the sequence area separates with the variance in so-called Formenraum.Der form area, in an educated manner from basis elements of defined stable form elements, reduces the complexity of the variants of the devices of the sequence area;
- the method opens the functional region over a variation of components of the form area;
- as components components of the form code (see below) are used;
- for the selection of the components certain criteria for choice are used for Vorselektion, which theoretical acceptance correspond or correspond to natural form analogues.

As modules for the parallel led variation (mutation) and selection so far only the nucleotides or amino acids are available as synthetic or enzymatically manageable components of a Polyiners for arranged coupling processes. The direct entrance to a functional surface structure of a polymer fails as aforementioned above in many to felling because of the problem of the large numbers of the variants of the sequence area.

Subject-matter of this evolutiven adjustment process is consisting the inset of modular components, the form code, of the form elements. The form code covers form elements, developed from elements of the sequence area. The form code, how it can being derived for example from natural polymers such as proteins, Polypeptiden or functional nucleic acids, codes stable form elements (secondary structures under determined outside conditions, possibly Tertiary period structural components containing). It is remarkable that very different sequences (primary structures) for very similar form elements can code.

With other words, in the form area very closely neighbouring elements can lie in the sequence area very far from each other far away (large Hamming distance). The same applies to the reverse case. In the sense according to invention evenly this property explains that already the replacement of form-moderately same sequences in the sequence area can mean a large step in the sense of a Vielfehlermutante. With the help of the addressed synthesis procedures according to invention this requirement is technically convertible. The production of the appropriate distributions succeeds by programmed synthesis. It is not, as in WHERE 92/18645 described to reach by incorrect Replikation in the sense of faulted PCR methods.

Fig. the term sequence area, form area and functional region describes 8. The analog regarded relationship of form area and sequence area applies to the relationship from form area and functional region that closely neighbouring, homologous elements can be from each other removed in the form area in the functional region far. As in fig. 8 schematically suggested, are determining for the function of a polymer geometry and the physikochemische topology and dynamics of the molecule surface, which steps with a second molecule into reciprocal effect. The underlying structure, defined from the form code, could be very different chemical nature. Similar functions in the function area explain themselves by similar boundary surface topologies.

Straight one regarding the relatively small molecule populations realizable in experiments, is it of crucial importance that the produced variation in the form area represents the possible function variety in the functional region in much more efficient way than for instance the variation in the sequence area.

The following descriptions of figure describe the invention by examples schematically more near.

The Fig. 1 concerns two einzelsträngige DNA and/or. RNA of molecules, those chemically or enzymatically (z. B. T4 RNA ligase) to be ligiert, whereby one of the molecules over fissile left one (z. B. Biotin Streptavidin) at solid phase immobilized is, while the other molecule is present freely in solution.

Stand to it today a whole set of solid phase materials (z. B. magnetic, surface-activated plastic balls) z order. This method permits the gradual structure of larger DNAs and/or. RNAs. After each Ligationsschritt not converted RNAs is away-washed and the Ligationsprodukte in the next Ligationsansatz, present on solid phase, is transferred. Favourable way is very simple the handling, in particular the purification of the respective Ligationsprodukte.

After termination of the last Ligation the product is used directly as Effektormolekül or translated in (in vitro) a

translation reaction first into the appropriate protein structure, which functions then as Effektormolekül.

The Fig. two completely doppelsträngige DNA of molecules, those concerns 2 chemically or enzymatically (z. B. To T4 DNA ligase) ?blunt end? to be ligiert, whereby one is imobilisiert over fissile left one at solid phase, while the other one is present freely in solution. In this way gradual larger doppelsträngige DNA molecules can be developed. The arranged Ligation is reached by different phosphorylation of the reaction partners. Module A and the last module are in such a way sketched that they in each case contain an interface for a restriction enzyme. This makes first of all the splitting off of the product from the solid phase and secondly the following, arranged Klonierung possible of the DNA (see also Fig. 5).

To Fig. 3: DNA molecules can in accordance with Fig. 2 to be ligiert likewise, if the molecule at a side, in solution, possesses a einzelsträngiges end, D. h. not completely doppelsträngig is present. This end does not stand in this way for the double-rank-specific Ligation, z. B. with T4 DNA ligase for the order. In combination with the phosphorylation strategies already mentioned (Fig. , in particular variant 1) the possibility results 2 of accomplishing the Ligation without unwanted by-products. The DNA molecule in solution can be so sketched that it before its einzelsträngigen end still the interface of a restriction enzyme preferably a Class the IIS of enzyme (z. B. AlwI) with recognition place in, the partial einzelsträngigen DNA piece which can be cut off) possesses. After the Ligation can the Ligationsprodukt at solid phase with the restriction enzyme be cut. In this way completely doppelsträngiges DNA molecule at solid phase develops. Alternatively the einzelsträngige end with a polymerase can be filled up to doubling rank or abverdaut with a Exonuklease.

To Fig. 4: Also restriction interfaces can develop (overlapping), as two doppelsträngige DNA molecules are ligiert with one another.

To Fig. 5: Completely or partial doppelsträngige DNA molecules can in accordance with Fig. 1-4 to be ligiert, even if mixtures of molecules (z. B. B, C, D) to be used. In this way develop mixtures of immobilized molecules, which correspond to different in each case combinations of the assigned components. At the end of the last Ligationsschrittes the entire DNA or a portion of it can by restriction enzymes, which cut within the Konstruktes, of the fixed phase abgespalten and if necessary. in a phage or a bacteria display system to be kloniert. In addition, the DNA can be exprimiert in a combined in-vitro-Transkriptions and lateral adjustment system.

To Fig. 6: On the basis of object module libraries can be produced peptides, protein domains and small proteins by coincidental combination of individual modules. According to a hierarchical method to the protein Design then also protein domains than components can become combined in a further stage. On each complexity stage mutations can be inserted, which - to change without the global structure - permit a fine tuning to the three-dimensional arrangement of chemical groups.

Fig. it describes 7 schematically that different proteins possess catalytic active amino acids in the active center despite more different, regarding the substrate homologous functions (Chymotrypsin/Trypsin) or despite similar spatial arrangement of the amino acids in the active center completely different reactions catalyzes (Trypsin/Elastase) can.

Fig. the connection of the terms sequence space form space functional region describes 8.

The sequence area is defined by the linear neighbourhood relations of the polymer devices of a polymer structure. Homologien describe similarities (in %) in the succession of the devices of a chemical material class. The more highly the degree of relationship of two sequences the smaller the spacing in the sequence area.

a) . . . AATAATGCGCAATATTAGGCCT. . .
b) . . . AATAAAAAGCAATATTAAGCCT. . .
c) . . . TTAGCTAGCGATGCGCGCCGGG. . .

For example the sequences A do not exhibit) and b) a substantial Homologie, while sequence C) any similarities with A) and b) shows.

The form area is defined by the ?spatial? neighbourhood relations of the polymers represented by it. The spacing of two sequences is certain by the degree of relationship of their structures. Homologie means here similarity of the total structures of polymers, which exist again out chemically linked devices. In the form area neighbouring molecules can lie in the sequence area quite far from each other far away and in reverse. [Analog S. o.: Structure A) 3 alpha Helices, structure b) 2 alpha Helices plus unstructured range with finalconstant, short Helix, C) antiparallel beta rabbet sheet from 4 sheets]. The function area is defined by the geometrical, dynamic and physical/chemical surface structure, which can step with a further molecule into specific reciprocal effect. Homologien describe similarities of the surface structure and the associated reciprocal effect characteristics.